

ABSTRACT OF THE DISCLOSURE

A hash function based data retrieval system for use with a lexicon database of a data processing system is disclosed. The system comprises a RAM, a disk memory, and a data file residing in the disk memory, wherein the data file contains stored data and is organized into nests. The system further comprises a hashing data structure residing in RAM, wherein the data structure is designed to occupy a fixed amount of memory independent of content of the data file. A data retrieval module is operable to identify a nest using a hash function that is based on parameters selected according to characteristics of the data file. The hash function is further designed to be optimized for content of the data file. The hash function is further designed to produce hash values based on the fixed amount of memory.